The Relation between Kin and Multi-level Selection

An Approach Using Causal Graphs

Samir Okasha

Abstract

Kin selection and multi-level selection are alternative approaches for studying the evolution of social behaviour, the relation between which has long been a source of controversy. Many recent theorists regard the two approaches as ultimately equivalent, on the grounds that gene frequency change can be correctly expressed using either. However this shows only that the two are *predictively* equivalent, not that they offer equally good causal representations of the evolutionary process. This paper articulates the notion of an 'adequate causal representation' using causal graphs, and then seeks to identify circumstances under which kin and multi-level selection do and do not satisfy the test of causal adequacy.

1 Introduction

The evolution of social behaviour has been analyzed in two different ways in the biology literature: the kin selection approach (KS) and the multi-level selection approach (MLS). The former emphasizes the genetic relatedness of social partners, while the latter emphasizes the relative strengths of within and between-group selection. The relation between the two approaches has been a source of controversy ever since it was first broached by Hamilton ([1975]). In earlier debates biologists tended to regard kin and multi-level (or group) selection as alternatives to each other (for example, Maynard Smith [1976], Dawkins [1982]). But many contemporary biologists maintain that KS and MLS models of social evolution are actually equivalent, on the grounds that gene frequency change can be correctly computed using either; and some have proposed translation manuals for going from one to the other.[1] Though dissenters from this equivalence claim can be found, the majority of social evolution theorists appear to subscribe to it.[2]

In this paper I examine the relation between KS and MLS from a new perspective. I grant (with one caveat) that the two are *predictively* equivalent, in that the evolutionary change in a social trait (or gene) can be correctly expressed by either (section 2). In this I follow the current orthodoxy. However, I argue that this predictive equivalence does not imply that the choice between the two approaches is a matter of subjective taste, as is often assumed. For evolutionary biology, like other sciences, aims to provide causal explanations; so a description of evolutionary change should ideally yield an adequate causal representation of the factors responsible for the change (section 3). My aim is to see whether this criterion can be used to discriminate

[1]The most explicit manual is found in Kerr and Godfrey-Smith ([2002]), who talk about 'individualist' and 'multi-level' perspectives. Queller ([1992]) and Dugaktin and Reeve ([1994]) also contain suggestions for relating KS to MLS. Lehmann et. al. ([2007]) provide a detailed account of how one particular model, originally formulated in MLS terms, may be re-construed in KS terms, *pace* its authors; Grafen ([2007]) does likewise.

[2]Grafen ([1984]) was one of the first to explicitly defend the equivalence thesis; other defenders include (Queller [1992]; Dugatkin and Reeve [1994]; Sober and Wilson [1998]; Kerr and Godfrey-Smith [2002]; Rousset [2004]; Lehmann et. al. [2007]; West, Griffin and Gardner [2007], [2008]; McIlreath and Boyd [2007]; Foster [2009]; Frank [1998], [2013]; Marshall [2011]; West and Gardner [2013]). Dissenters from the equivalence thesis include (van Veelen [2009]; Hölldobler and Wilson [2009]; Traulsen [2010]; Nowak, Tarnita and Wilson [2010]; Wilson [2012]).

between the KS and MLS approaches to social evolution.

The notion of an 'adequate causal representation' obviously needs unpacking. I start by illustrating the notion with a pair of examples in which, intuitively, only one of KS or MLS counts as causally adequate (section 4). An explicit account of what it means for a description of evolutionary change to be 'causally adequate' is then developed, using causal graphs (section 5). This enables us to identify cases in which KS is preferable to MLS and cases in which the reverse is true. The concluding discussion (section 6) traces the implications of the theory developed here, relates the theory to extant work in the literature, and speculates about empirical examples which might exemplify the abstract causal relations depicted in the diagrams.

2 The KS and MLS Approaches

Determining the relation between KS and MLS is complicated by the fact that neither has a canonical formulation that all parties agree on. To circumvent this problem, I consider a simplified model of social evolution that makes minimal assumptions, and use it to derive maximally general formulations of both KS and MLS, that permits their explicit comparison.

Consider a population of haploid individuals living in groups of the same size, within which social interactions occur (Figure 1). Generations are non-overlapping. An allele at a particular locus codes for a social behaviour. We define $p_i = 1$ if the i^{th} individual has the allele, and $p_i = 0$ otherwise; we refer to p_i as the 'p-value' of individual i.[3] Note that the index i ranges over all individuals in the global population, irrespective of grouping. The population-wide frequency of the allele is $\overline{p}$, i.e. the average of all the individual p-values.

The reproductive output ('fitness') of individual i, defined as the total number of successful gametes it contributes to the next generation, is denoted w_i. The average fitness in the population is $\overline{w}$. Mutation is assumed absent. Under these assumptions, the change in allele frequency over a single generation is given by:

$$\Delta\overline{p} = \frac{Cov(w_i, p_i)}{\overline{w}} \tag{1}$$

[3]This terminology is adapted from Grafen ([1985]) who uses 'p-score' to denote the frequency of an allele within an individual or linear combinations thereof.

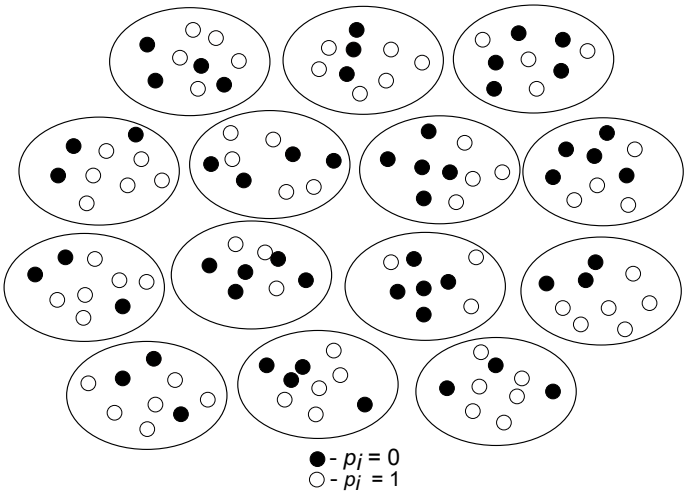


Figure 1: **Individuals in a group-structured population**

as is well-known (Price [1970], Robertson [1966]).

Equation (1) tells us that the allele, and thus the social behaviour that it codes for, will spread so long as $Cov(w_i, p_i) > 0$, i.e. there is a positive covariance between an individual's fitness and its p-value. This is highly intuitive, formalizing the core neo-Darwinian idea that genes associated with higher individual fitness will increase in frequency.

Equation (1) is always true but not always useful, as it does not discriminate between different *reasons* why an allele may covary positively with individual fitness. One possible reason is that the allele directly enhances the fitness of its bearers; another is that the allele causes its bearers to act in a way that enhances the fitness of others, for example, by coding for an altruistic action towards relatives. These two possibilities are quite different in terms of the phenotypes they will give to rise to, but equation (1) holds true in either case. Indicative of this is that although we specified that our allele codes for a social behaviour, this fact plays no role in the derivation of (1).

To produce a more useful analysis of social evolution, the covariance term in (1) can be decomposed into elements that are (hopefully) more biologically meaningful. In effect, the KS and MLS approaches constitute alternative ways of performing this decomposition. We consider them in turn.

2.1 The MLS decomposition

Since the individuals are nested in groups, the covariance term $Cov(w_i, p_i)$ can be written as the sum of 'between-group' and 'within-group' components, following Price ([1972]). We let p_{jk} and w_{jk} denote, respectively, the p-value and fitness of the j^{th} individual within the k^{th} group; we let P_k and W_k denote, respectively, the average p-value and average fitness of the individuals in the k^{th} group (referred to henceforth as 'group p-value' and 'group fitness'.) It then follows that:

$$Cov(w_i, p_i) = \overbrace{Cov(W_k, P_k)}^{\text{between-group}} + \overbrace{E_k[Cov(w_{jk}, p_{jk})]}^{\text{within-group}} \qquad (2)$$

The first RHS term of equation (2) is the covariance between group p-value and group fitness; while the second RHS term is the average across groups of the *within-group* covariance between individual p-value and fitness. So the former measures the extent of 'between-group' selection, while the latter measures the extent of 'within-group' selection. Substituting (2) into (1) then gives:

$$\overline{w}\Delta\overline{p} = \overbrace{Cov(W_k, P_k)}^{\text{between-group}} + \overbrace{E_k[Cov(w_{jk}, p_{jk})]}^{\text{within-group}} \qquad (3)$$

Equation (3) is more useful than (1) for understanding social evolution, and was employed by Price ([1972]) and Hamilton ([1975]) for that purpose. To see why, suppose that the allele codes for an 'altruistic' behaviour which is individually costly but group-beneficial. In that case, the first RHS term of equation (3) will be positive—groups with a greater p-value will have higher group fitness—while the second RHS term will be negative–within each group, individual p-value will correlate negatively with fitness. Thus equation (3) captures the idea that within-group and between-group selection may 'pull' in different directions, an idea central to much theorizing in the MLS tradition. It is easy to see why Hamilton ([1975]) described equation (3) as effecting a 'formal separation of the levels of selection' (p. 333), and why modern advocates of MLS often use it as their central organizing tool.[4]

[4]Works that accord a key role to equation (3) include (Queller [1992], Sober and Wilson [1998], Michod [1999], Okasha [2006], McIlreath and Boyd [2007] and Bowles and Gintis [2011]). Note that equation (3) describes multi-level selection of the MLS1 variety, in the terminology of Damuth and Heisler ([1988]).

From equation (3), it is straightforward to extract a condition for the spread of the allele, namely $Cov(W_k, P_k) > E_k[Cov(w_{jk}, p_{jk})]$, i.e. the between-group component of selection should exceed the within-group. Where this inequality is satisfied, the allele will increase in frequency in the next generation. More generally, equation (3) teaches us that the direction of selection on a social trait depends crucially on the balance between the within-group and between-group genetic variance.

Finally, note that equation (3) rests on minimal assumptions. To derive (3) from (1), we assumed only that the individuals in the population were nested into non-overlapping groups of equal size.[5] Nothing was said about the nature of these groups, their persistence, mode of reproduction, functional organization, nor the nature of inter-group competition. Thus equation (3) is really an overarching model-schema which captures the core MLS logic, and subsumes more specific models of evolution in structured populations. Let us refer to equation (3) as the 'MLS decomposition' of the overall evolutionary change.

In taking equation (3) as a formalization of MLS, I am following the orthodoxy among social evolutionists. However some authors, for example, Goodnight, Schwartz and Stevens ([1992]), advocate the alternative 'contextual analysis' approach to MLS, which partitions the total change up differently. On that approach, the difference between MLS and KS is considerably reduced, since contextual analysis and KS are in fact very similar, conceptually and formally, as noted by Goodnight ([2013]). I do not discuss contextual analysis further here (see Okasha [2006]), but simply assume, with the bulk of the recent social evolution literature, that MLS is appropriately captured by equation (3).

2.2 The KS decomposition

The key idea of the KS approach is that there are two distinct routes by which an allele can increase its frequency in a population: enhancing the reproductive output of its bearer (the direct route), or enhancing the reproductive output of other carriers of the allele (the indirect route). The roots of this idea are present in Darwin ([1859]) and Haldane ([1932]), but it was Hamilton ([1964]) who first made it

[5]The equal size assumption is for convenience only. To deal with unequally sized groups, we simply need to weight the covariance and expectation terms in equation (3) by group size, as per Price ([1972]).

explicit, with his concept of 'inclusive fitness' and his famous $rb > c$ rule for the spread of altruism.

The KS approach can be applied to our abstract model to yield another decomposition of the total change, using a well-known technique.[6] Since the allele codes for a social action, an individual's fitness w_i depends on both its own p-value (p_i), and on the average p-value of its social partners (p_i'), i.e. the other group members with whom it interacts. We can thus write w_i as a linear regression on p_i and p_i':

$$w_i = \alpha + \beta_{wp.p'} p_i + \beta_{wp'.p} p_i' + e_i \tag{4}$$

where α is 'baseline' fitness; $\beta_{wp.p'}$ is the partial regression of individual fitness on individual p-value, controlling for social partners' p-value; $\beta_{wp'.p}$ is the partial regression of individual fitness on social partners' p-value, controlling for individual p-value, and e_i is the residual. Note that equation (4) does not presume that the true dependence of w_i on p_i and p_i' is linear; the regression equation can always be fitted, thanks to the residual.

Equation (4) can then be substituted into equation (1), which after simplification yields:

$$\overline{w}\Delta\overline{p} = \big(\beta_{wp.p'} + \beta_{wp'.p}\beta_{p'p}\big)Var(p) \tag{5}$$

where $\beta_{p'p}$ is the regression of social partners' p-value on individual p-value, and thus measures the genetic similarity among social partners.

To put equation (5) in more familiar form, we can then re-label the $\beta_{wp.p'}$ and $\beta_{wp'.p}$ coefficients as '$-c$' and 'b' respectively, and the $\beta_{p'p}$ coefficient as 'r', to give:

$$\overline{w}\Delta\overline{p} = (-c + rb)Var(p) \tag{6}$$

Equation (6) yields Hamilton's rule in its standard form, i.e. the allele will spread so long as $rb > c$. It is important to see why the re-labelling used to get from (5) to (6) is justifiable. Take the 'r' term first. Though Hamilton originally defined 'r' in genealogical terms, it is a familiar point that what really matters is the genetic similarity of social interactants, not their genealogical relatedness; indeed $\beta_{p'p}$ is nowadays a fairly standard definition of 'r' in the kin selection literature.[7]

[6] This technique is used and discussed by (Queller [1992], [2011]; Frank [1998], Gardner, West and Wild [2011]; Birch [forthcoming]), among others.

[7] For discussion of this point see (Hamilton [1975]; Grafen [1985]; Queller [1992], [2011]; Sober and Wilson [1998]; McIlreath and Boyd [2007]; Okasha [2008]; Birch [forthcoming]).

Turning to '$-c$', this refers to the cost, i.e. reduction in personal fitness, incurred by an individual performing the social action. A natural measure of this cost is the difference in personal fitness, on average, between two individuals with the same social partners (i.e. same value of p'_i), one of whom has the allele ($p_i = 1$), so performs the social action, the other of whom does not ($p_i = 0$); and this is exactly how $\beta_{wp.p'}$ is defined.

Turning to 'b', this is often defined as the benefit, i.e. increase in personal fitness, that an individual performing a social action confers on a recipient. Here we define 'b' slightly differently, as the average benefit that an individual *receives* from its social partners, rather than confers on them, which is what the $\beta_{wp.p'}$ coefficient measures. This means that equation (6) is a 'neighbour-modulated' rather than an 'inclusive fitness' version of KS theory.[8]

To facilitate explicit comparison with the MLS decomposition, it helps to re-write (6) as:

$$\overline{w}\Delta\overline{p} \;=\; \overbrace{(-c)Var(p)}^{\text{direct effect}} \;+\; \overbrace{rbVar(p)}^{\text{indirect effect}} \qquad (7)$$

Equation (7) partitions the total change into 'direct' and 'indirect' components, capturing the idea that the overall allele frequency change depends on the direct effect of the allele on its bearers and on the indirect effect on relatives. Let us refer to (7) as the 'KS decomposition'. Note that the KS decomposition deals only with individual fitness—the expressions for group fitness (W_k) and group p-value (P_k) do not feature in it; this is the key point of contrast with the MLS decomposition.

The KS decomposition when derived this way is quite general, since (7) follows from (1) with no new assumptions. Many models of kin selection, including Hamilton's original, make restrictive assumptions, for example, that selection is weak, or that costs and benefits are additive. However these assumptions are not needed to derive the $rb > c$ rule, *if* we are happy to define 'r', 'b' and 'c' as above, as Gardner, West and Wild ([2010]) stress.[9]

[8]For discussion of this distinction, see (Taylor and Frank [1996]; Rousset [2004]; Frank [1998]; Wenseleers, Gardner and Foster [2010]; Taylor, Wild and Gardner [2007]; Birch [unpublished]).

[9]Some critics, for example, Nowak, Tarnita and Wilson ([2011]), argue that defining 'r', 'b' and 'c' as partial regression coefficients deprives them of biological meaning, as they

3 Equivalence and Causality

The foregoing analysis helps explain why the MLS and KS approaches are often regarded as equivalent. In any group-structured population, the total evolutionary change can be decomposed using either equation (3) or (7). Moreover, it is easy to see that the KS criterion for spread of a pro-social trait ($rb > c$), will be satisfied if and only if the MLS criterion (between-group > within-group) is satisfied. Thus the two approaches are *predictively* equivalent. Gene frequency change can be computed in two ways: by determining the magnitude of the between and within-group components, or the direct and indirect effects; both methods will always give the same answer.

In one respect the KS approach is more general. For in deriving (7) we did not use the fact that the individuals are nested into non-overlapping groups. All we assumed was that an individual's fitness is affected by its 'social partners'. These social partners were stipulated to be fellow group members, but this played no role in the derivation. It follows that the KS decomposition can apply to a population where individuals interact with neighbours but there are no discrete groups. However the MLS decomposition requires non-overlapping groups. Though this difference is important, it has been discussed elsewhere and is orthogonal to our main concerns here, so we leave it aside.[10]

Granting the predictive equivalence of KS and MLS, it does not follow that the choice between the two is a subjective or aesthetic matter, as many authors imply. For the aim of evolutionary analysis is not solely to predict changes in gene frequency, but also to provide causal explanations of them (*inter alia*). So ideally we want a description of evolution to provide insight into the causal factors responsible for the evolutionary change in question, in addition to computing the correct answer.

The notion of causality rarely features explicitly in evolutionary models; but it is usually assumed that at least some of the dependencies between a model's variables reflect causal relationships in nature. For example if a population genetics model assigns different viabili-

become functions of population-wide gene frequencies. This is an important objection to the 'generalized' version of Hamilton's rule given in equation (7); however it raises issues orthogonal to the ones treated in this paper.

[10]See (Maynard Smith [1976]; Okasha [2005], [2006]; Godfrey-Smith [2008]; Frank [2013]) for discussion of this point.

ties to different genotypes, it is taken for granted that the viability difference is meant to be caused by the genotypic difference (via some phenotypic intermediary). This causal assumption does not feature explicitly in the mathematics, but it is essential if the model is to be of use in understanding real-world systems.

Turning to KS and MLS, are there circumstances in which one provides a more adequate causal representation of social evolution than the other? Someone might answer 'no' on the following grounds: "suppose a gene for a social trait increases in frequency, i.e. $\Delta\overline{p} > 0$, and we wish to explain why this is. One possible explanation is that $rb > c$; another is that between-group selection exceeds within-group. But these two inequalities hold in exactly the same circumstances, therefore KS and MLS are explanatorily and not just predictively equivalent."

However this argument misses the point (though helps to clarify the issue). For when we say that $\Delta\overline{p} > 0$ because the inequality $rb > c$ is satisfied (for example), this is not strictly speaking a causal attribution. The satisfaction of the inequality does not literally *cause* $\Delta\overline{p}$ to be non-zero, or to assume the particular value that it does. So while it is true that the MLS and KS inequalities are satisfied in exactly the same circumstances, it does not follow that the respective decompositions—equations (3) and (7)—constitute equally good causal representations of the evolutionary process.

The idea of a description of evolutionary change being 'causally adequate', or 'true to the causal facts', has often been appealed to in the literature on levels of selection, but has never been made fully precise.[11] When the total evolutionary change is written using a statistical partition, as above, a natural explication of this notion suggests itself: the statistical associations between variables should reflect direct causal influences in the world.[12] This idea is spelled out in detail in section 5, with the aid of causal graphs.

To fix ideas, and to connect with the literature, I turn first to two 'problem cases'. In each, one of the two approaches (KS and MLS) is causally adequate but the other is causally misleading, intuitively. These cases will help illustrate the notion of causal adequacy.

[11] See in particular (Sober and Lewontin [1982]; Sober [1984]; Okasha [2006]; Godfrey-Smith and Kerr [2013]; Frank [2013]; Birch [forthcoming]).

[12] This appears to correspond to what Okasha ([2006]), Frank ([2013]) and Birch ([forthcoming]) refer to as a 'causal decomposition' of the total change, though none of these works makes the notion precise.

4 Two problem cases

4.1 The non-social trait case

The first case was introduced by Sober ([1984]) and Nunney ([1985]); it is also discussed by (Heisler and Damuth [1987]; Okasha [2004a], [2004b], [2006]; Godfrey-Smith and Kerr [2013]). Consider again a group-structured population as in section 2. Suppose that the allele in question codes for a non-social trait, for example, a physiological property; so an individual's fitness depends only on its own p-value, not on the p-value of its fellow group members. Suppose also that most of the genetic variance is between group, for whatever reason. Thus the group p-values vary widely, but within any group the p-values of individuals tend to be similar. Let us call this the 'non-social trait' case.

When the MLS decomposition (equation (3)) is applied to this case, it yields the result that the between-group component of evolutionary change $Cov(W_k, P_k)$ is large, while the within-group component $E_k[Cov(w_{jk}, p_{jk}]$ is very small. So group-level selection is largely responsible for the evolutionary change.

Intuitively this is an undesirable result. For it seems clear that the trait for which the gene codes is an individual adaptation, that spreads because it benefits individuals not groups. Groups in which the trait is common have higher group fitness, but this is a side-effect of the fact that individuals with the trait are fitter than individuals without (see Okasha [2006]). Indicative of this is that the group structure makes no difference to the evolutionary outcome; $\Delta\overline{p}$ would be the same even if the individuals didn't live in groups. So while $Cov(W_k, P_k)$ is large, this does not reflect a direct causal influence of group p-value P_k on group fitness W_k.

This point is related to G.C. Williams' ([1966]) distinction between 'group adaptation' and 'fortuitous group benefit'.[13] Williams argued that just because some feature of a group benefits the group doesn't mean that it evolved for that reason; it could be an incidental side-effect of selection acting on individuals. But the MLS decomposition is insensitive to Williams' distinction; for it does not distinguish between cases where $Cov(W_k, P_k)$ reflects a direct causal influence of P_k on W_k and cases where it is 'caused from below'. In the latter, such as the non-social trait case, the MLS decomposition is causally inadequate,

[13]Williams often uses the term 'biotic adaptation' in place of 'group adaptation'.

inviting the mistaken inference that the trait spreads because it is group-beneficial.

Defenders of MLS have proposed two different solutions to this problem. Sober ([2011]) argues that the MLS decomposition should only be applied where 'groups' exist in some substantial sense. If we define groups as entities within which fitness-affecting interactions take place (with respect to the trait of interest), as per Sober and Wilson ([1998]), then in the non-social trait case there are no groups in the first place.[14] A different approach is taken by Heisler and Damuth ([1987]), Goodnight, Schwartz and Stevens ([1992]) and Okasha ([2006]), who argue that the MLS decomposition should be replaced by 'contextual analysis' , which tries to isolate a genuine group effect on individual fitness, thus distinguishing 'real' group selection from cases of 'causation from below'.

Here is not the place to discuss these proposed 'fixes' to the MLS approach. But note that the KS approach deals neatly with the non-social trait case.[15] To see this, recall equation (7), which partitions the total change into direct and indirect components. In the non-social trait case the indirect component will be zero; this is because an individual's fitness depends only on its own p-value, so the b coefficient in Hamilton's rule will be zero. Thus the KS decomposition *is* causally adequate in this example, correctly suggesting that the trait evolves because it directly benefits the individuals that have it, not their social partners or their groups.

4.2 Genotypic selection with meiotic drive

The second case is an example devised originally by Okasha ([2004a]), and discussed in Okasha ([2006]) and Godfrey-Smith and Kerr ([2013]). The example involves 'frameshifting' the multi-level apparatus downwards, by treating diploid organisms as groups of genes.[16] Thus consider a structured population in which the individuals are alleles and

[14]See Okasha ([2004a], [2011]) for discussion of this attempt to deal with the non-social trait case.

[15]The KS approach is in fact closely related to contextual analysis. The only difference is that in the linear regression model (4), in place of the p-value of the individual's social partners (p_i'), contextual analysis uses as a regressor the p-value of the group to which the individual belongs. See Okasha ([2004b]) and Goodnight ([2013]) for further discussion of the relation between KS and contextual analysis.

[16]It is a familiar point that diploid population genetics can be regarded as a type of multi-level selection by frameshifting in this way; for discussion of this point see (Wilson

the groups are diploid genotypes, i.e. groups of size n=2. For simplicity, consider a single locus with two alleles, A and B, and thus three genotypes, AA, AB and BB. There are fitness differences between the genotypes: $w_{AA} = 16, w_{AB} = 12, w_{BB} = 8$. In addition, segregation in the AB heterozygote is distorted in favour of A in the ratio 2:1, i.e. of the 12 successful gametes that an AB organism produces, 8 contain A and 4 contain B.[17] The net evolutionary outcome depends on both the genotypic fitness differences and the meiotic drive. Let us call this the 'genotypic-selection-with-drive' case.

This case is naturally analysed using the MLS approach. The fitness differences between genotypes constitutes between-group selection, while the meiotic drive constitutes within-group selection. Thus we can define $p_i = 1$ if the i^{th} allele is A, $p_i = 0$ otherwise, and analyze $\Delta\overline{p}$ using the MLS decomposition (equation (3)). Both of the right-hand side terms will be non-zero, reflecting the fact that both within and between-group selection are playing a role. Although in this example the two levels of selection pull in the same direction—favouring the A allele—the MLS framework still makes good sense here, since it is clear that two distinct selective forces are at work.

The KS approach, by contrast, yields an untoward result, namely that the direct effect is solely responsible for the evolutionary change, i.e. the indirect effect in equation (7) is zero. For notice that the fitness of an individual A allele is the same whether it is in an AA or an AB organism; in either case it leaves 8 copies in the next generation. So given the fitness scheme and the specific pattern of meiotic drive posited, any allele's fitness depends only on its own p-value, not on the p-value of its social partner. Therefore the b coefficient in Hamilton's rule equals zero, so the indirect effect is zero. Formally this is isomorphic to the non-social trait case, frameshifted down a level in the biological hierarchy.

In this case the KS approach fails the test of causal adequacy, as it invites the mistaken inference that the A allele spreads solely because it is advantageous to the individual allele, and not because it is advantageous to the group (i.e. host organism). Whereas in reality, the spread of the A allele *is* partly due to the fact that an organism's fitness is causally influenced by the proportion of A alleles it contains. (Were there no fitness differences between genotypes, the

([1997]; Okasha [2004c]; Gardner, West and Barton [2007]).

[17] The key feature of the example is that segregation in the heterozygote is distorted in the ratio $w_{AA} : w_{BB}$.

A allele would spread at a slower rate.)

The genotypic-selection-with-drive case shows that sometimes the MLS approach is causally adequate while the KS approach is not. There is also a more general moral, namely that isomorphism of fitness structures does not necessarily imply isomorphism of the causal facts giving rise to those fitness structures. It is striking that although the genotypic-selection-with-drive and the non-social trait cases are formally identical, given the particular parameter values chosen, the MLS approach provides causal insight in the former but not the latter case, and vice-versa for the KS approach.

To make progress, we need to clarify the notion of causal adequacy, then to find a general account of when KS and MLS do and do not yield causally adequate descriptions of evolution in structured populations.

5 Causal adequacy: a graphical approach

5.1 The basic idea

The basic idea to be defended is simple. A decomposition of evolutionary change is causally adequate if and only if the statistical associations that appear in the components reflect direct causal influences between the variables in question.[18] So for example, if there is a correlation between group p-value P_k and group fitness W_k but no causal influence of the former on the latter, as in the non-social trait case, then the MLS decomposition is causally inadequate.

There are two ways in which a decomposition can fail to be causally adequate: (i) it contains statistical associations between variables which do not causally influence each other; or (ii) it contains statistical associations between variables which only indirectly causally influence each other. (The meaning of 'indirect' will become clear.) Intuitively (i) is a more serious form of causal inadequacy than (ii). These ideas are best illustrated by using causal graphs, or path diagrams.

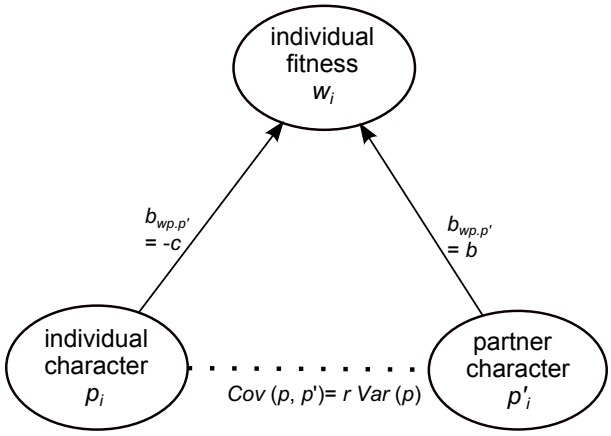
We continue to use the structured population model of section 2. Every individual i in the global population has a p-value p_i, a fitness w_i, and an average social partners' p-value p'_i. For some purposes we will need to explicitly model the (average) fitness of an individual's

[18] It is taken as read here that one of the variables is fitness (of either individuals or groups). Thus for equation (7) to be causally adequate, it is not required that the r term, i.e. $\beta_{p'p}$, should reflect a direct causal influence of p on p', obviously.

social partners, which we denote by w_i'. (Note that w_i' does not occur explicitly in either the MLS or KS decompositions.)

Suppose firstly that the gene codes for a social trait, so w_i is causally influenced by both p_i and p_i', which are themselves correlated. This is reflected in Figure 2, in which arrows denote causal influences and dotted lines denote correlations. The strength of the causal influences are measured by (unstandardized) path coefficients, i.e. partial regressions, and of the correlations by simple covariances. Note that w_i and p_i are connected by two distinct pathways, one direct and the other indirect.

Figure 2 denotes a situation in which the KS decomposition is causally adequate, since the '$-c$' and 'b' terms of equation (7) reflect direct causal influences between p_i and w_i, and p_i' and w_i, respectively. The overall evolutionary change $\Delta\overline{p}$ is given by $Cov(w_i, p_i)$, as usual; this covariance is the net result of the two distinct causal paths from p_i to w_i, which correspond to the $-c$ and b components of Hamilton's rule. Any case of social evolution involving causal dependencies as in Figure 2 is naturally analyzed using the KS approach.



$$\overline{w}\Delta\overline{p} = Cov(w_i, p_i) = (-c)Var(p) + (rb)Var(p)$$

Figure 2: **Case where KS is causally adequate**

The value of $Cov(w_i, p_i)$ can be read directly off Figure 2 using the rules of path analysis, yielding the KS decomposition (equation (7)). These rules, first laid out by Sewall Wright, tell us how to compute the

covariance between any two variables on an acyclic causal graph; the basic idea is simply to find each distinct path from the dependent to the independent variable, compute the product of the path coefficients on each path, then sum over all the paths. In Figure 2, there are two paths from w_i to p_i, one direct and one direct. The products of the coefficients on each path are $(-c)$ and $(rb)Var(p)$ respectively. (However since the variables are unstandardized, the $(-c)$ coefficient on the direct path needs to be multiplied by $Var(p)$).[19] Summing over the two paths, we can thus deduce the value of $Cov(w_i, p_i)$, as in equation (7).

5.2 Graphs with individual and group variables

Note that the key variables of the MLS decomposition, group p-value P_k and group fitness W_k, do not appear in Figure 2. To facilitate comparison between MLS and KS, we need to add these variables to the causal graph. A change of notation is needed to do this. Instead of P_k and W_k, which range over groups, we write P_i and W_i, which range over individuals; these denote the average p-value and average fitness of the group to which the i^{th} individual belongs. So P_i and W_i are in effect relational properties of individual i. Similarly, in place of p_{jk} and w_{jk}, which denote the p-value and fitness of the j^{th} individual in the k^{th} group, we may write p_{ji} and w_{ji}, which denote the p-value and fitness of the j^{th} individual in the group to which the i^{th} individual belongs. This notational change does not affect the MLS decomposition above, since $Cov(W_k, P_k) = Cov(W_i, P_i)$, and $E_k[Cov(w_{jk}, p_{jk})] = E_i[Cov(w_{ji}, p_{ji})]$.

To depict individual and group variables on the same diagram, we need to attend to the logical relations between them. This raises some subtleties. Since a group's p-value is defined as the average p-value of its constituent individuals, and similarly for group fitness, the following relationships must hold between P_i, p_i and p'_i, and between W_i, w_i and w'_i:

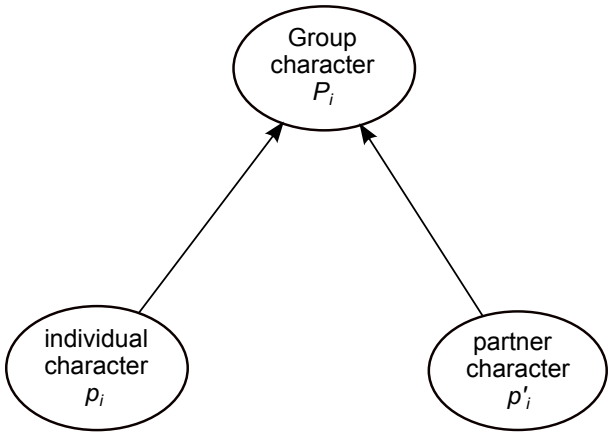
$$W_i = (w_i + (n-1)w'_i)/n \tag{8}$$

$$P_i = (p_i + (n-1)p'_i)/n \tag{9}$$

where n is group size.

[19] See Heise ([1975]) for a good discussion of the path analysis rules for unstandardized variables.

Equations (8) and (9) express mathematical dependencies; any one of the three variables $\{P_i, p_i, p'_i\}$ is a linear function of the other two, and similarly for $\{W_i, w_i, w'_i\}$. Despite this, in the case of equation (9) at least, there is an intuitive asymmetry. It is individuals that have p-values (genes) in the first instance, a group only has a p-value (gene frequency) in virtue of the p-values of its constituent individuals; group p-value is a 'mere aggregation' of individual p-values. So metaphysically, P_i is determined by p_i and p'_i, not vice-versa. It is natural to depict this asymmetry graphically by drawing arrows from p_i and p'_i to P_i, as in Figure 3 below.



$$P_i = (p_i + (n-1)p'_i)/n$$

Figure 3: **Relation between individual, partner and group characters**

It might be objected that the dependencies depicted in Figure 3 are not really causal. Strictly speaking surely P_i supervenes on, or is metaphysically determined by, p_i and p'_i, rather than being caused by them? Arguably this is true (though some authors, for example, Searle ([1992]), hold that a supervenience relationship can itself be causal). However the real issue is whether a relation of metaphysical dependence can be depicted as if it were causal without violating the principles of causal modelling. I think the answer is 'yes', though there is one complication.[20]

[20]This question does not seem to have raised in the causal modelling literature, though

In the causal modelling literature, the meaning of the arrows is usually explained in terms of hypothetical interventions (for example, Pearl ([2000])). An arrow between X and Y means that if one were to 'intervene on', i.e. experimentally manipulate, variable X, while keeping everything else fixed, then Y would change too. The arrows in Figure 3 do satisfy this interpretation. Intervening on either p_i or p_i' while keeping the other fixed will necessarily change the value of P_i. In this respect, the relations between p_i and P_i and p_i' and P_i are akin to causal relations, even if strictly speaking they are relations of metaphysical determination.

The complication comes when we consider intervening on P_i itself, i.e. changing the group p-value. For it is impossible to change P_i while holding p_i and p_i' fixed, given the supervenience of the former on the latter.[21] One might argue that for this very reason, it is impossible for variables such as P_i to have causal effects of their own, and thus that in any causal graph containing the variables P_i, p_i and p_i', no causal arrows should come out of P_i. However this is a controversial metaphysical issue which it is better not to pre-judge, so the following convention is adopted here.[22] In a causal graph in which one variable supervenes on others, when we consider hypothetically intervening on the supevenient variable we do *not* hold fixed the variables on which it supervenes, but rather alter them to preserve consistency. *Modulo* this convention, causal arrows going out of supervenient variables, if any, can be understood in the usual way.

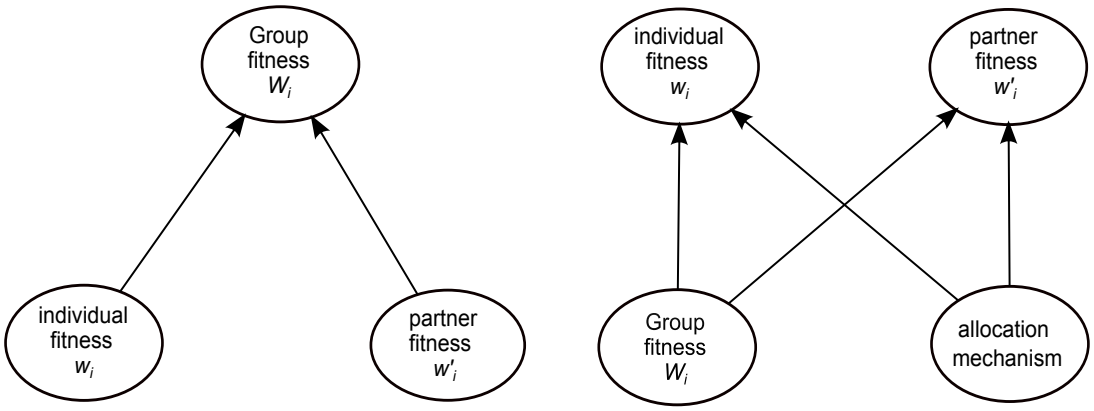
What about the relation between W_i, w_i and w_i'? In some cases it will parallel the relation between P_i, p_i and p_i'; but this need not be so, for fitness is a different sort of property to p-value. If the group is a mere collection of individuals each of which is able to survive and reproduce independently, then intuitively fitness is a property of individuals in the first instance, and group fitness W_i is derivative. In such cases then W_i is metaphysically determined by w_i and w_i', just as for P_i, p_i and p_i'. However in other cases it is the whole group that has a fitness in the first instance, which is then shared out among the individuals in the group by an allocation mechanism; see below. In such cases an arrow runs from W_i to w_i (and to w_i'), rather than vice-

see Shipley ([2000], pp.43-5) on how to deal with 'conserved quantities' in a causal model.

[21] This is related to what are called 'fat hand' interventions in the causal modelling literature; see Scheines ([2006]).

[22] The issue here is closely connected to the discussion of 'causal exclusion' in the philosophy of mind; see Kim ([1998]).

versa, i.e. the fitness of any individual is determined by the fitness of its group and the allocation mechanism. These two possibilities are depicted in Figure 4. Note that equation (8) holds true in both cases.



4a **4b**

$$W_i = (w_i + (n-1)w'_i)/n$$

Figure 4: **Relation between individual, partner and group fitness**

To better understand Figure 4b, consider an example. Suppose a group is engaged in cooperative hunting, and gets a particular quantity of meat on a given day—the 'group payoff'—which is shared out among the individuals in the group (fairly or otherwise). We can think of the group and individual payoffs as proxies for group and individual fitness. There is a mathematical relation between the individual payoffs and the group payoff: the former add up to the latter. But intuitively, it is the group payoff plus the allocation rule that determines the individual payoffs, not vice-versa. The explanation of why any individual received the particular payoff they did, is that the whole group received a particular payoff, a certain portion of which was allocated to that individual. Conversely, the explanation of why the group received the payoff it did is *not* that each of the individuals received the payoff that they did. So the group payoff is primary and the individual payoffs are derivative. Similarly, in certain cases group fitness will be primary and individual fitness derivative, as in figure 4b.

Where Figure 4b applies, and in the co-operative hunting example above, it may be difficult to decide whether the arrow from W_i to w_i should be interpreted as causation or metaphysical determination. Is an individual's fitness caused by its group's fitness (along with the allocation mechanism), or does it supervene on it? Fortunately we do not need to resolve this issue, as nothing of substance depends on it. What is important is the contrast between Figures 4a and 4b.

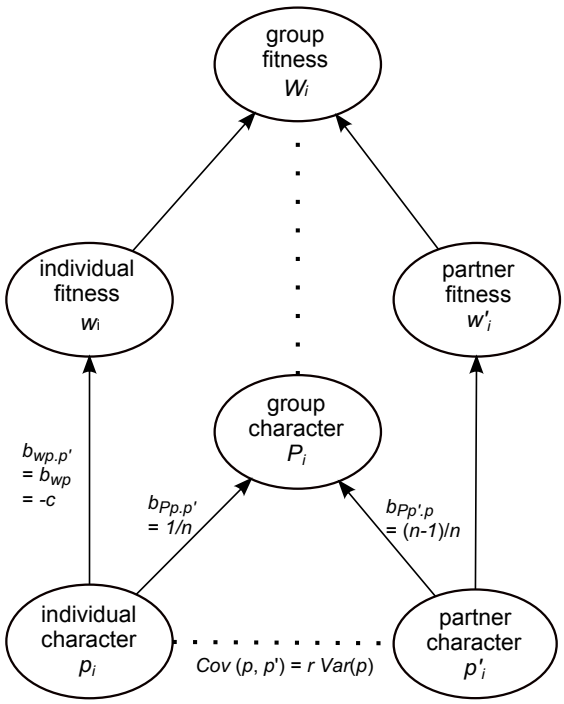
To summarize, it *is* possible to incorporate variables that are mathematically related, such as individual and group variables, on a causal graph. Firstly we need to determine the direction of metaphysical determination between the variables. (This may involve modal intuitions, empirical knowledge of the system being modeled, or both). We then draw arrows that reflect this metaphysical determination, which can be interpreted in the usual way. Secondly, we adoption the convention that when intervening on a supervenient variable, we do not hold fixed the variables on which it supervenes, but rather adjust them to preserve consistency.

5.3 Cases where KS is causally adequate

Now consider the non-social trait case, in which an individual's fitness depends only on its own p-value, with no contribution from social partners. This is depicted in Figure 5, where the only arrow going in to w_i is from p_i. Similarly, the fitness of individual i's social partners, w_i', depends only on p_i'. In this case the group fitness W_i is determined by w_i and w_i', just as in Figure 4a: what leads a group to have the fitness value it does is precisely that its constituent individuals have the fitness values that they do. Hence the arrows from w_i and w_i' to W_i.

Note that in Figure 5, group p-value P_i correlates with but does not causally influence group fitness W_i. The reason for the correlation is clear: these two variables are joint effects of a common cause (p_i). If we increase the p_i value of individual i, this will increase P_i; it will also increase w_i and thus W_i. By contrast, changing p_i' has no effect on w_i, indicating the absence of social partners' contribution to individual fitness. The absence of a causal arrow between P_i and W_i means that the MLS decomposition is causally inadequate: the term $Cov(W_i, P_i)$ does not reflect a causal influence. By contrast, the KS decomposition is causally adequate.

Suppose we modify Figure 5 by adding causal arrows from p_i' to w_i



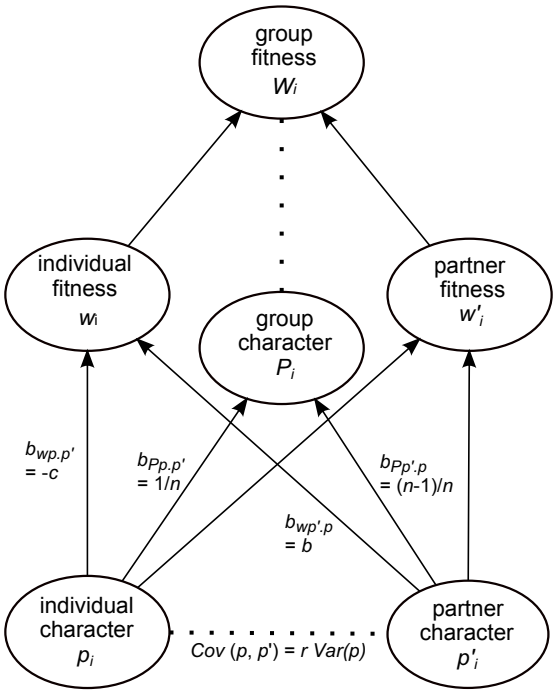
$$\overline{w}\Delta\overline{p} = Cov(w_i, p_i) = b_{wp.p'}Var(p) = (-c)Var(p)$$

Figure 5: **Non-social trait case**

and from p_i to w'_i, yielding Figure 6. Thus we no longer have a non-social trait, since an individual's fitness is now causally influenced by social partners. However in Figure 6 the MLS approach is still causally inadequate. For there is still no causal arrow between group p-value P_i and group fitness W_i—they are joint effects of a common cause. So the $Cov(W_i, P_i)$ term of equation (3) still does not reflect a causal influence of P_i on W_i. Conversely, the KS approach is causally adequate—for both the b and c terms of equation (7) reflect direct causal influences, as the figure shows. This shows that the presence of fitness-affecting interactions between social partners is not sufficient for the MLS approach to be causally adequate; the group-level covariance may still be a non-causal side effect.

What biological cases does Figure 6 apply to? The key features are the presence of social interactions between individuals within groups, but the absence of a causal influence of group p-value P_i on either

individual or group fitness. This means that the causal explanation of any individual's fitness can be given in terms that refer only to individual-level properties, i.e. p_i and p'_i. This will be true if there is no group-level functional organization, and the groups exhibit no emergent properties of their own, i.e. the only group properties are aggregations of individual properties, such as P_i. Many mammalian social groups, for example, buffalo herds and baboon troops, arguably satisfy these conditions. Extensive social interactions take place within such groups, but the groups are not functionally integrated in the way that eusocial insect colonies are, for example.



$$\overline{w}\Delta\overline{p} = Cov(w_i, p_i) = (-c)Var(p) + (rb)Var(p)$$

Figure 6: **Social trait case where KS is causally adequate, MLS not**

Suppose instead that we modify Figure 5 by adding a causal arrow from P_i to both w_i and w'_i, yielding Figure 7. In this causal scheme, each individual's fitness is causally influenced by its own p-value and also by the p-value of the group to which it belongs, and thus only

indirectly by social partners' p-value. This causal scheme will apply if a group's p-value gives rise to further group-level 'emergent properties' (not modeled here), which in turn affect individual fitness. If so, then the causal explanation of any individual's fitness will involve group as well as individual properties, unlike in Figure 6. However, group fitness is still derivative: a group only has a fitness value in virtue of the fitnesses of its constituent individuals.

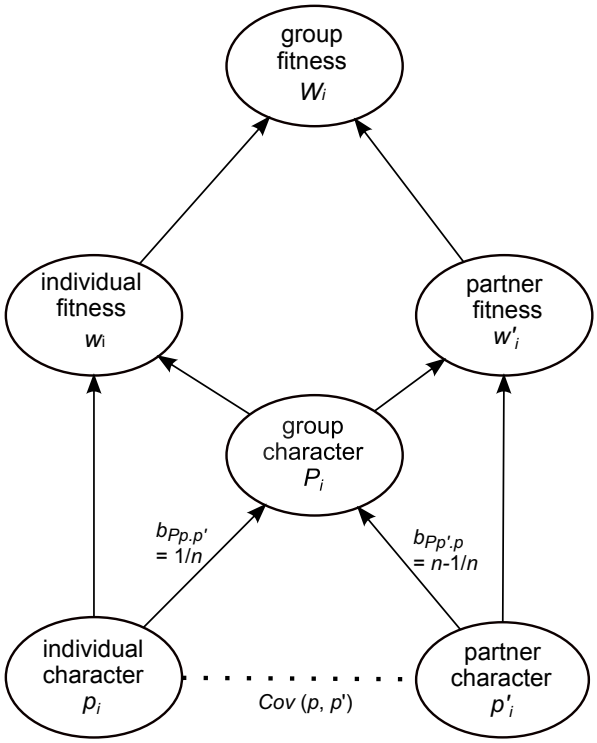


Figure 7: **Social trait case where neither KS nor MLS is causally adequate**

Figure 7 represents an interesting case, in that neither KS nor MLS is causally adequate. KS is inadequate for the second of the two reasons mentioned above: although p_i and p'_i do causally influence w_i, the latter does so only indirectly, via its effect on group p-value

P_i. So the b and c terms of the KS decomposition do not reflect direct causal influences. Similarly, MLS is inadequate because although group p-value P_i does influence group fitness W_i, it does so indirectly, via its effect on w_i and w_i'. So the $Cov(W_i, P_i)$ term of the MLS decomposition does not reflect a direct causal influence. The criterion of causal adequacy offers no basis for preferring one approach to the other in this case, though note that in both cases the inadequacy is of the second, less serious sort.[23]

5.4 Cases where MLS is causally adequate

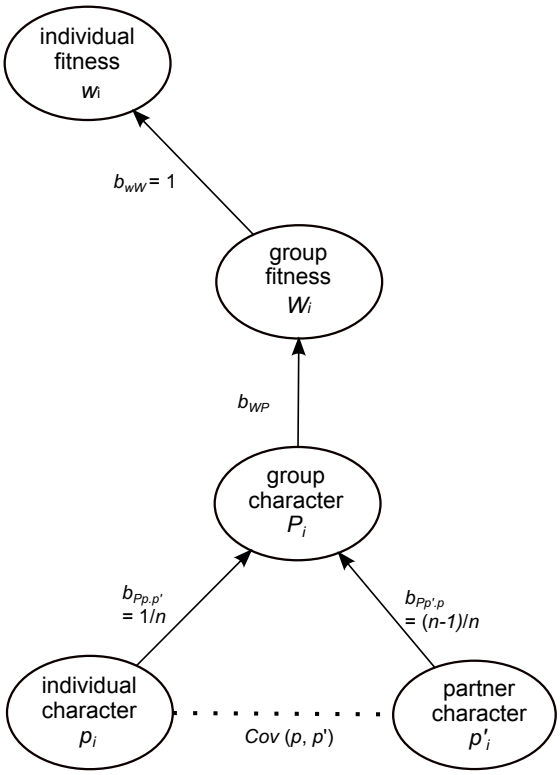
To depict a case where MLS is causally adequate, let us frameshift downwards again as in section 4.2; so the 'individuals' are alleles and the 'groups' are diploid organisms of size n=2. As before we consider a single locus with two alleles A and B, and thus three genotypes AA, AB and BB. We define p_i=1 if the i^{th} allele is an A, 0 otherwise; so $\overline{p}$ is the population-wide frequency of the A allele. Assume firstly that meiosis is fair, i.e. within-group selection is absent; but there are fitness differences between genotypes. This scenario, which I call 'Genotypic selection with fair meiosis' is depicted in Figure 8 below.

Note that in Figure 8, a causal arrow leads from group fitness W_i to individual fitness w_i (as in Figure 4b above). This is because it is a group (i.e. diploid organism) that has a fitness value in the first instance, i.e. it contributes a certain number of successful gametes to the next generation, depending on its genotype. This gametic output is shared equally among group members (i.e. the two alleles), owing to the fairness of meiosis, so the fitness of an individual allele (w_i) is simply equal to the fitness of its group (W_i), which is itself causally determined by the group's p-value (P_i), which in turn is determined by individual p-value (p_i) and partner p-value (p_i').

There are two key features of Figure 8. Firstly, group fitness W_i is primary and individual fitness w_i derivative. The fitness of any individual (i.e allele) is determined by the fitness, or total gametic output, of the group (i.e. diploid organism) to which it belongs. So an arrow runs from W_i to w_i, not vice-versa.[24] Secondly, group fitness

[23] Arguably, a causally adequate decomposition of $\Delta\overline{p}$ in Figure 7 would be provided by 'contextual analysis', which constitutes a kind of hybrid of the KS and MLS approaches; see footnote 15 above.

[24] An arrow also runs from W_i to partner fitness w_i' in this example, but the variable w_i' is omitted from the graph for simplicity.



$$\overline{w}\Delta\overline{p} = Cov(w_i, p_i) = Cov(W_i, P_i) = b_{WP}Var(P)$$

Figure 8: **Genotypic selection with fair meiosis**

W_i is caused by group p-value (P_i), and thus only indirectly by p_i and p'_i. The reason for this is clear. The full causal explanation of why an organism has a particular fitness value (W_i) would refer to its phenotype, which depends in a complex way on its genotype (P_i), both of which are whole-organism properties. These is no reason to hold that properties of alleles, i.e. p and p_i, are doing the 'real' causal work, nor therefore that the link between P_i and W_i is a mere non-causal correlation.

Figure 8 depicts a situation in which MLS is causally adequate. Within-group selection is absent, so the overall evolutionary change $\Delta\overline{p}$ is proportional to $Cov(W_i, P_i)$, i.e. the covariance between group p-value and group fitness, which is by definition equal to $b_{WP}.Var(P_i)$.

Since there is a direct causal influence of P_i on W_i, the strength of which is measured by b_{WP}, it follows that the MLS decomposition is causally adequate.

By contrast, the KS decomposition is inadequate as applied to Figure 8. For again, the causal influences of p_i on w_i, and of p'_i on w_i, are both indirect, mediated by group fitness W_i. To see this, pick an individual with p_i=0, i.e. a B allele, and consider hypothetically altering its p-value to p_i=1, i.e. mutating it into an A. This will alter the genotype of the allele's host organism, i.e. P_i will increase by 1/2, which in turn will alter the organism's fitness W_i, thus indirectly affecting the individual allele's own fitness w_i. Similarly, altering p'_i affects w_i only indirectly, via affecting W_i. Thus the KS decomposition is causally inadequate; for the b and c components of equation (7) reflect highly indirect causal influences.

To appreciate this point from another angle, contrast Figures 6 and 8. In both cases, p_i and p'_i are the ultimate causal determinants of individual fitness w_i; and the KS decomposition can be applied to both, if desired. However in Figure 6, p_i and p'_i directly cause w_i, while in Figure 8 the causation is indirect, mediated by group p-value P_i and group fitness W_i. So Figure 6 depicts a case where KS is causally adequate but MLS is not; while Figure 8 depicts a case where MLS is causally adequate but KS is not.

The contrast between Figures 6 and 8 highlights that whether a causal influence between variables counts as direct or indirect depends on which other variables we are considering. In describing the influence of p_i on w_i in Figure 6 as 'direct', this does not mean that there are no causal intermediaries at all between p_i and w_i, but just that none of the other variables in our model (for example, P_i or W_i) is such an intermediary. Thus the notion of directness is model-relative. This is a corollary of a point stressed in the causal modeling literature, namely that a choice must always be made about which variables to include in a causal graph and which to omit (see Pearl [2000]; Spirtes, Glymour and Scheines [2001]).

It is interesting to consider whether the causal structure depicted in Figure 8 might apply to any cases in which the 'individuals' are whole organisms and the 'groups' are social groups. The crucial feature of Figure 8 is that an arrow leads from group fitness W_i to individual fitness w_i, rather than the other way around, and that this is the sole determinant of w_i. This means that the group as a whole attains a given fitness value, depending on its group properties; and any indi-

vidual's fitness value is a causal consequence of its group having the fitness value that it does. Plausibly, this might apply to certain co-operatively breeding groups, or eusocial insect colonies, in which any individual's reproductive success is heavily dependent on the group, and some mechanism exists for suppressing reproductive competition within the group. We return to this issue in the next section.

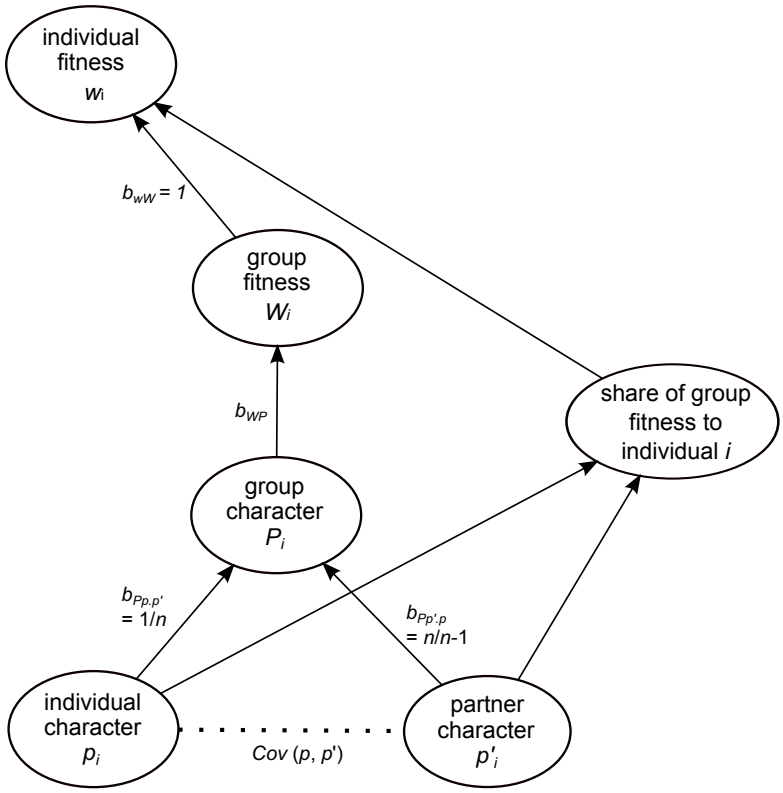
Lastly, consider the second 'problem case' from section 4.2: genotypic-selection-with-drive. Suppose that the A allele is a meiotic driver. This case is depicted in Figure 9, which differs from Figure 8 in that an individual allele's fitness w_i now depends not just on group fitness W_i, but also on the *share* of the group's fitness that accrues to it, i.e. the 'allocation mechanism'. Since meiosis is not fair, this share is not constant for all i. The share of group fitness that accrues to individual i is jointly determined by p_i and p'_i. (This share may be measured by the deviation of an individual's fitness from the mean fitness of its group.) If $p_i = p'_i$, i.e. if the individual is in an AA or BB homozygote, then the group fitness is shared equally; while if $p_i \neq p'_i$, i.e. the individual is in an AB heterozygote, then the group fitness is unequally shared, due to meiotic drive.

The causal structure in Figure 9 highlights the two different routes by which an individual allele can increase its fitness : (i) by increasing the total fitness of its group, or (ii) by increasing its share of the group's fitness. (If fair meiosis is in place then route (ii) is closed off, and we are back to Figure 8). These two causal pathways correspond to the between-group and within-group components of the MLS decomposition.

Figure 9 depicts a situation where MLS is causally adequate, in that P_i directly causes W_i, so the between-group component of the MLS decomposition, $Cov(W_i, P_i)$, reflects a direct causal influence. Unlike in Figure 8, there is also a component of within-group selection, i.e. meiotic drive; so the total change $\Delta\overline{p}$ is given by the sum of the within-group and between-group components, as in equation (3). This means that the second causal pathway from p_i to w_i, that goes via the variable 'share of group fitness to individual i', corresponds to the within-group term of the MLS decomposition.[25]

Note that in Figure 9, two causal pathways run from partner p-value p'_i to individual fitness w_i. To understand this, consider an

[25]Note that this within-group term, i.e. $E_k[Cov(w_{jk}, p_{jk})]$(or equivalently $E_i[Cov(w_{ji}, p_{ji})]$), though corresponding to a specific causal pathway on Figure 9, cannot be directly read off the figure, since the w_{ji} and p_{ji} variables do not explicitly feature in Figure 9.



$$\overline{w}\Delta\overline{p} = \quad Cov(W_i, P_i) \quad + \quad E_i([Cov(w_{ji}, p_{ji})])$$

Figure 9: **Genotypic selection with meiotic drive**

individual A allele in an AA homozygote. This individual has p_i=1, p'_i=1 and P_i=1. Consider intervening to set p'_i=0, i.e. mutating the partner into a B. This intevention has two effects: it converts the host organism into an AB heterozygote, i.e. makes $P_i = \frac{1}{2}$, thus altering the group fitness W_i; and it also increases the share of the group fitness that accrues to individual i. In section 4.2, the chosen parameter values ensured that these two effects exactly cancel. (These values were: $w_{AA} = 16, w_{AB} = 12, w_{BB} = 8$, with segregation in AB distorted 2:1 in favour of A.) With these parameter values, reducing p'_i from 1 to 0 decreases group fitness by 4, but increases individual i's

share of the group fitness by 4. So it *appears* that individual fitness w_i is causally uninfluenced by partner p-value p'_i. However, this is an illusion: p'_i and w_i are connected by *two* causal pathways, and the causal influences transmitted along the two pathways exactly cancel out. Thus w_i appears to depend only on p_i.

We can now see clearly the difference between the non-social trait case (Figure 5), and the genotypic-selection-with-drive case (Figure 9). In the former, w_i is caused solely by p_i, as there is no causal path from p'_i to w_i. So altering p'_i has no effect on w_i. In the latter, there are two causal paths from p'_i to w_i; but altering p'_i has no net effect on w_i because, given the particular parameter values chosen, the two causal influences exactly cancel out. So the fitness values in the two cases are isomorphic, but the causal structures that give rise to them are quite different. This in turn explains why the KS approach is causally adequate in the non-social trait case but the MLS approach is not; and vice-versa in the genotypic-selection-with-drive case.

Interestingly, the circumstance where a causal influence along one path exactly cancels that along another is well-known in the literature on causal modeling. Pearl ([2000]) refers to this as 'instability', while Spirtes, Glymour and Scheines ([2001]) call it 'unfaithfulness'. This phenomenon complicates the task of inferring causality from correlational data, for it implies that the joint probability distribution on a set of observed variables will be a misleading guide to the underlying causal structure. This is precisely what we see in the genotypic-selection-with-drive case, and is precisely why the KS approach as applied to this case is misleading. The above authors emphasize, however, that instability / unfaithfulness is likely to be a rare circumstance in practice, as it only arises for specific parameter values. This point also applies in the genotypic-selection-with-drive case, as the canceling effect only occurs given the specific combination of genotype fitness values and meiotic drive posited.

Again, it is interesting to consider whether Figure 9 might apply to organisms in social groups, rather than genes in genomes. The crucial features are that group fitness determines individual fitness, but the group's fitness is unequally divided; so individuals can enhance their fitness either by boosting group fitness, or by boosting their share of it. Conceivably this could apply to social groups with co-operative breeding and/or reproductive division of labour, where individuals can only reproduce as part of their group, but in which reproductive competition among group members is not fully suppressed. In such

scenarios, the MLS approach is likely to be causally adequate.

6 Discussion

The foregoing analysis has a number of implications for the debate between KS and MLS. Firstly, the widespread view that the choice between the two approaches is a matter of modelling preference, or computational convenience, is untenable. The predictive equivalence of KS and MLS does not imply that they constitute equally adequate causal representations of evolutionary change. Secondly, claims for the global superiority of one of the two approaches are untenable.[26] Given our analysis of causal adequacy, it seems likely that empirical cases exist where only one of the two approaches counts as causally adequate.

It is interesting to note that in cases where the KS approach fails the test of causal adequacy (Sigures 7, 8 and 9), this is always for the second, less serious reason, i.e. the b and c terms of equation (7) reflect indirect causal influences. By contrast in some cases, the MLS approach fails to be causally adequate for the more serious reason, namely the 'between-group' term of equation (3) reflects a correlation between variables that do not causally influence each other at all, but are joint effects of a common cause (Figures 5 and 6). This difference relates to the fact that in all the causal graphs above, p_i and p_i' are the ultimate causes of individual fitness w_i, thus the KS decomposition can only fail to be adequate because p_i and p_i' cause w indirectly.

This consideration highlights the fact that the KS approach is reductionistic, or individualistic. Such an approach must always be possible, since the evolutionary change of interest, $\Delta\overline{p}$, is the change in the population mean of an individual property p_i.[27] $\Delta\overline{p}$ is given by the global covariance between p_i and fitness w_i, as equation (1) tells us, which is the net result of two causal paths: from p_i to w_i, and from p_i' to w_i. So the overall evolutionary change is always 'ultimately' attributable to these two causal paths. However it does not follow that the KS approach is always superior. Where the causal paths from p_i

[26]Such claims have been made for KS by West, Griffin and Gardner ([2008]), Bourke ([2011]) and others; and for MLS by Sober and Wilson ([1998]), Nowak, Tarnita and Wilson ([2010]) and others.

[27]This reflects the fact that the type of multi-level selection in question is MLS1 rather than MLS2, in the terminology of Heisler and Damuth ([1987]).

and p_i' to w_i are mediated by group p-value and group fitness, as in Figures 8 and 9, the MLS approach offers more insight. This is an instance of the general moral that reductionistic explanations in science are not always superior.

Our analysis identifies one issue, in particular, as critical to the choice between KS and MLS: whether an arrow leads from individual fitness w_i to group fitness W_i, or vice-versa. In the non-social trait case, and arguably in some social trait cases, the arrow runs from w_i to W_i. In such cases, fitness and reproduction are properties of individuals, and a group only has a fitness value as a consequence of its constituent individuals having their fitness values. In other cases, including the 'genes in genomes' cases depicted in Figures 8 and 9 and arguably in certain 'organisms in groups' cases too, fitness pertains to the whole group in the first instance, and any individual has a fitness value only as a consequence of its group having a fitness value.

The distinction between w_i determining W_i, and the converse, is clear at the abstract level of causal graphs; but it is not easy to say what biological features determine which side a given case falls on. One relevant feature is whether individuals are capable of surviving and reproducing independently of their group. If they are, then it is natural to regard a group's fitness as an aggregation of the individuals' fitnesses, even if there are extensive social interactions, so the arrow runs from w_i to W_i. But if they are not, then the group is the 'primary' bearer of fitness, and individual fitness is derivative. By the same token, if there is extensive reproductive division-of-labour within a group, with some individuals specialized in somatic and others in reproductive tasks, as in some eusocial insect colonies, then the group is naturally treated as the primary bearer of fitness. This conceptualization is standard among at least some social insect biologists (for example, Seeley [1997]; Hölldobler and Wilson [2009]).

Where group fitness is primary, then an allocation mechanism must determine the share of the group's fitness that each individual receives. One possibility is that a mechanism exists that suppresses within-group competition; if the mechanism works perfectly then all individuals in the group receive equal shares, so $W_i = w_i$ for all i, as in Figure 8. Numerous such mechanisms exist in nature (Frank [2003]). Fair meiosis is one such example; others include policing of worker reproduction in social insects (Ratnieks and Reeve [1992]), enforced monogamy in human groups (Alexander [1987]), punishment of selfish actions in human groups (Bowles and Gintis [2011]), and restricting

cell-lineages' access to the germ-line in multicellular organisms (Buss [1987]). These mechanisms all tend to suppress the variation in individual fitness within groups.

In cases where group fitness is primary and the suppression mechanism works near-perfectly, it is common to treat the whole group as the adapted unit rather than as an aggregation of lower-level units. Biologists think like this when they treat a multi-celled organism as an adapted unit rather than a mere aggregation of cell-lineages, a eukaryotic cell as an adapted unit rather than a mere aggregation of prokaryotic units, and a diploid genome as an adapted unit rather than a mere aggregation of alleles. The theory presented here supports this conceptualization, since in such cases (depicted in Figure 8), the MLS approach comes out as causally adequate, and the entirety of the evolutionary change is attributable to between-group selection. But even if the suppression mechanism is not perfect, the MLS approach may still be causally adequate as long as an arrow runs from W_i to w_i, as in Figure 9; there is then a component of within-group selection. The important contrast is with cases where individual fitness is primary, i.e. an arrow runs from w_i to W_i, in which the MLS approach is causally inadequate.

In certain empirical cases, it may be indeterminate which causal graph fits the case best, and thus indeterminate whether IF or MLS is causally adequate. Possible examples include cellular slime molds, social bacteria colonies, certain *Volvox* colonies, and certain marine invertebrate colonies; in these cases, it is unclear whether W_i metaphysically determines w_i or vice-versa. But this is not a problem; indeed some such indeterminacy is only to be expected, given the literature on evolutionary transitions.[28] This literature teaches us that social interaction among free-living individuals living in groups, for example, cells in cell-groups, was likely a precursor stage in the evolution of new higher-level biological units, for example, multi-celled organisms, in which the original individuals have lost the ability to survive and reproduce independently and are mere parts of the whole. The theory presented here implies that in the early transitional stages, only KS would have been causally adequate, as group fitness would have been derivative from individual fitness, while in the latter stages, only MLS would have been causally adequate. Since evolutionary transitions are inevitably gradual, a zone of indeterminacy is thus to be expected.

[28]See in particular (Buss [1987]; Maynard Smith and Szathmáry [1995]; Michod [1999]; Queller [1997]; Godfrey-Smith [2009]; Bourke [2011]; Calcott and Sterelny [2011].)

6.1 Relation to Previous Work

It is worth relating the theory presented above to previous work. In recent incarnations of the 'KS versus MLS' debate in the biology literature, a number of authors have argued in favour of one approach and against the other. These arguments are of two main sorts. Their protagonists either reject the predictive equivalence of KS and MLS, or else accept that equivalence but advance 'pragmatic' considerations in favour of one of them. The former is an option because to prove the predictive equivalence of KS and MLS, the 'r', 'b' and 'c' terms must be defined in a particular way; theorists such as Nowak, Tarnita and Wilson ([2011]) regard these definitions as lacking biological meaning, and thus dispute the claim that KS covers all cases of social evolution. The latter is an option because predictively equivalent descriptions may differ in respect of computational tractability, or ease of empirical application; it is on this basis that authors such as West, Griffin and Gardner ([2007], [2008]) and Frank ([2013]) favour KS over MLS, for example.

The theory presented here also seeks to discriminate between KS and MLS, but in a different way. My approach has been to grant, at least for the sake of argument, the assumptions necessary to establish the predictive equivalence of KS and MLS, but to emphasize another dimension—causal adequacy—on which to compare the two approaches. Unlike pragmatic factors, which are partly subjective, the question of whether a given description of evolutionary change is causally adequate is fully objective, so long as it is clear which causal graph is applicable.

Our theory prompts comparison with recent work by Godfrey-Smith and Kerr ([2013]), who contrast 'contextual' and 'collective' ways of modeling evolution in structured populations, building on their earlier work. (This contrast corresponds closely to the KS versus MLS distinction.) They lay out an abstract theory for choosing between these two modelling approaches, and test it against a variety of problem cases (including the 'non-social trait' and 'genotypic-selection-with-drive' cases of section 4.2 above.) Godfrey-Smith and Kerr mention 'causal adequacy', but do not explicitly depict the hypothesized causal relations between the state-variables of their evolving system. Instead they devise a criterion, the 'near-variant' test, for deciding whether a given case should be modeled in 'contextual' or 'collective' terms (given that both options are always available). In

effect, the near-variant test involves modifying the fitness structure slightly, then seeing whether the modification is more simply accommodated by the contextual or collective approach, by asking how many parameters must be changed in each case.

The 'near-variant test' is interesting, and delivers intuitively plausible answers. However the test has a rather pragmatic flavour, so it is unclear whether it can be more than a heuristic guide to the true causal structure of the evolutionary process. The approach taken here is different, in that the hypothesized causal relations between variables are explicitly represented on a graph, and an explicit criterion of causal adequacy is proposed; it is then possible to specify exactly what the true causal graph must be like, in order for a candidate statistical partition of the evolutionary change to be causally adequate. Logically this is a quite different procedure to that of Godfrey-Smith and Kerr, so it is striking that their test and our theory yield similar verdicts in at least some cases. Further work on the relation between our theory and the near-variant test might be useful.

Recent work by Gardner and Grafen ([2009]) studies the conditions under which groups may be considered 'adapted units', using Grafen's 'formal Darwinism' framework. They argue that clonal groups, and groups in which within-group reproductive competition is completely suppressed, are the only situations in which group adaptationism is valid. (Okasha and Paternotte ([2013]) re-visited Gardner and Grafen's analysis and found that the argument only holds in the repression-of-competition case.) Though these papers are about group *adaptation* rather than group (or multi-level) *selection*, which is the focus here, it is natural to expect a connection; and there is one. For complete suppression of within-group competition corresponds to Figure 8 above; in this case the MLS approach is causally adequate and within-group selection is null. (By contrast, if groups are clonal this does *not* imply the causal adequacy of MLS, since it is possible that we are dealing with a non-social trait case, as in Figure 5.) Thus the circumstances under which 'group adaptationism' is valid, according to Gardner and Grafen's analysis as amended by Okasha and Paternotte, are precisely those in which the MLS approach is causally adequate and within-group selection is null, according to the theory presented here.

Lastly, Michod and collaborators have outlined a theoretical framework for studying 'evolutionary transitions in individuality' and applied it to the evolution of multi-cellularity, using $Volvox$ as a model

system.[29] A key idea in their framework is that during a transition, the fitness of a group eventually becomes 'decoupled' from that of its constituent individuals, so the group becomes the new Darwinian individual. One way this can occur is through reproductive division-of-labour, i.e. germ/soma specialization; when complete, such specialization means that lower-level individuals cease to be autonomous evolutionary units. Another is through the evolution of policing mechanisms to suppress within-group conflict (cf. Clarke ([2013])).

Michod does not explicitly relate his theory to the KS versus MLS issue, but it is natural to suggest a link. In the early transitional stages, prior to fitness decoupling, individuals engage in social interactions; at this stage individual fitness is primary and group fitness derivative, so KS will be causally adequate. Once fitness decoupling has occurred, group fitness becomes primary, and the MLS approach becomes causally adequate. If this is correct, then Michod's notion of 'fitness decoupling' offers a way of fleshing out the abstract graphical approach developed above, as it provides a criterion for helping decide which causal graph applies to which empirical cases. Prior to fitness decoupling, the relevant causal structure will look something like Figure 6, while after fitness decoupling it will look like Figure 8 or 9. Note, however, that this connection with Michod's work is a plausible conjecture, not something that has been formally shown.

7 Conclusion

The opposition between the kin and multi-level selection approaches to social evolution has long polarized evolutionary biologists. In recent discussions theorists have increasingly come to regard the two approaches as equivalent, on the grounds that a correct expression for allele frequency change in structured populations can be written using either approach. However this establishes only that KS and MLS are *predictively* equivalent, not that they constitute equally adequate causal descriptions of the evolutionary process. The bulk of this paper has been concerned with fleshing out, in a precise way, the meaning of 'causal adequacy', and using it to determine when the KS and MLS decompositions of the total evolutionary change count as

[29] See (Michod [2005], [2006], [2007]); Michod and Roze [1999]; Michod *et al.* [2006]). For discussion of Michod's notion of 'fitness decoupling', see Okasha ([2006]), [2009]) and Bossert, Qi and Weymark ([2013]).

causally adequate.

The theory presented here does not pretend to be complete. Even if the theory is broadly on the right lines, a number of important issues remain. One is to get clearer about the implicit metaphysical commitments involved in the multi-level causal graphs above, which contain both individual and group variables. Another is to address the epistemological problem of how we can tell which causal graph is correct, in any given case. Finally, further work to map actual empirical cases onto the abstract framework developed here would be useful; the cases discussed in the evolutionary transitions literature are an obvious starting point. These all represent potential avenues for future work.

Finally, the need for the above theory stems from a striking disjunction between the scientific explanations that evolutionary biologists aim to give, which are causal, and the formal models of the evolutionary process that they develop, in which causal concepts do not feature explicitly. This disjunction is not unique to evolutionary biology, less still to social evolution theory, but is a quite general problem; it stems from the fact that a mathematical framework adequate for describing causal relations has only been developed fairly recently (Pearl [2000]; Spirtes, Glymour and Scheines [2001]). Importing this framework into biology, and philosophy of biology, is an important task for the future; pioneering first steps were taken by Shipley ([2000]), but much remains to be done.

Funding

European Research Council Seventh Framework Program (FP7/2007–2013), ERC Grant agreement (295449).

Acknowledgments

Thanks to Ellen Clarke, Chris Clarke, Jonathan Birch, Steve Frank, Elliott Sober, Andy Gardner, James Marshall, Kevin Foster, Jun Otsuka, Omri Tal, Cedric Paternotte, Ken Binmore, David Crawford, Johannes Martens, Peter Godfrey-Smith and two anonymous referees for discussion and comments.

Samir Okasha

Department of Philosophy,
University of Bristol,
Bristol BS6 6JL,
United Kingdom
Samir.Okasha@bristol.ac.uk

References

[1] Alexander, R. D. [1987]: *The Biology of Moral Systems*, New York: Aldine.

[2] Birch, J. [unpublished]: *Kin Selection: a Philosophical Analysis*, PhD thesis submitted to the University of Cambridge.

[3] Birch, J. [forthcoming]: 'Hamilton's Rule and its Discontents', *British Journal for the Philosophy of Science*

[4] Bossert, W., Qi, C. X. and Weymark, J. A. [2013]: 'Extensive Social Choice and the Measurement of Group Fitness in Biological Hierarchies', *Biology and Philosophy*, **28**, pp. 75–98.

[5] Bourke, A. F. G. [2011]: *Principles of Social Evolution*, Oxford: Oxford University Press.

[6] Bowles, S. and Gintis, H. [2011]: *A Cooperative Species*, Princeton, NJ: Princeton University Press.

[7] Buss, L. [1987]: *The Evolution of Individuality*, Princeton, NJ: Princeton University Press.

[8] Calcott, B. and Sterelny, K. [2011]: *The Major Transitions in Evolution Revisited*, Cambridge, MA: MIT Press.

[9] Clarke, E. [2013]: 'The Multiple Realisability of Biological Individuals', *Journal of Philosophy*, CX, **8**, pp.413–435.

[10] Clarke E. and Okasha S. [2013]: 'Species and Organisms: What Are the Problems?', in F. Bouchard and P. Huneman (*eds*), *From Groups to Individuals: Evolution and Emerging Individuality*, Cambridge, MA: MIT Press, pp. 62–84.

[11] Damuth J. and Heisler I. L. [1988]: 'Alternative Formulations of Multi-level Selection', *Biology and Philosophy*, **3**, pp. 407–30.

[12] Darwin, C. [1859]: *On the Origin of Species by Means of Natural Selection*, London: John Murray.

[13] Dawkins, R. [1982]: *The Extended Phenotype*, Oxford: Oxford University Press.

[14] Dugatkin L. A. and Reeve H. K. [1994]: 'Behavioural Ecology and Levels of Selection: Dissolving the Group Selection Controversy', *Advances in the Study of Behavior*, **23**, pp. 101–33.

[15] Foster, K. R. [2009]: 'A Defense of Sociobiology', *Cold Spring Harbour Symposium in Quantitative Biology*, **74**, pp. 403–418.

[16] Frank, S. A. [1998]: *Foundations of Social Evolution*, Princeton, NJ: Princeton University Press.

[17] Frank, S. A. [2003]: 'Repression of Competition and the Evolution of Cooperation', *Evolution*, **57**, pp. 693–705.

[18] Frank, S. A. [2013]: 'Natural Selection. VIII. History and Interpretation of Kin Selection Theory', *Journal of Evolutionary Biology*, **26**, pp. 1151–84.

[19] Gardner, A. and Grafen, A. [2009]: 'Capturing the Superorganism: a Formal Theory of Group Adaptation', *Journal of Evolutionary Biology*, **22**, pp. 659–671.

[20] Gardner A, West, S. A. and Barton, N. H. [2007]: 'The Relation between Multilocus Population Genetics and Social Evolution Theory', *American Naturalist*, **169**, pp. 207–26.

[21] Gardner, A., West, S. A. and Wild, G. [2011]: 'The Genetical Theory of Kin Selection', *Journal of Evolutionary Biology*, **24**, pp. 1020–43.

[22] Godfrey-Smith, P. [2008]: 'Varieties of Population Structure and the Levels of Selection', *British Journal for the Philosophy of Science*, **59**, pp. 25–50.

[23] Godfrey-Smith, P. [2009]: *Darwinian Populations and Natural Selection*, Oxford: Oxford University Press.

[24] Godfrey-Smith, P. and Kerr, B. [2013]: 'Gestalt-Switching and the Evolutionary Transitions', *British Journal for the Philosophy of Science*, **64**, pp. 205–22.

[25] Goodnight, C. J., Schwartz, J. M. and Stevens, L. [1992]: 'Contextual Analysis of Models of Group Selection, Soft Selection, Hard Selection, and the Evolution of Altruism', *American Naturalist*, **140**, pp. 743–761.

[26] Goodnight, C. J. [2013]: 'On Multi-Level Selection and Kin Selection: Contextual Analysis Meets Direct Fitness', *Evolution*, **67**, pp. 1539–48.

[27] Grafen, A. [1984]: 'Natural Selection, Kin Selection and Group Selection', in J. R. Krebs and N. B. Davies (*eds*), 1984, *Behavioural Ecology, 2nd edition*, Oxford: Blackwell Scientific Publications, pp. 62–84.

[28] Grafen, A. [1985]: 'A Geometric View of Relatedness', *Oxford Surveys in Evolutionary Biology*, **2**, pp. 28–89.

[29] Grafen, A. [2007]: 'Detecting Kin Selection At Work Using Inclusive fitness', *Proceedings of the Royal Society B*, **274**, pp. 713–19.

[30] Haldane, J. B. S. [1932]: *The Causes of Evolution*, London: Longmans, Green and Co.

[31] Hamilton, W. D. [1964]: 'The Genetical Evolution of Social Behaviour, I and II', *Journal of Theoretical Biology*, **7**, pp. 1–16, 17–52.

[32] Hamilton, W. D. [1975]: 'Innate Social Aptitudes in Man: an Approach From Evolutionary Genetics', in R. L. Fox (*ed*), 1975, *Biosocial Anthropology*, New York: John Wiley, pp. 115–32, reprinted in Hamilton, W. D. [1996], pp. 315–352.

[33] Hamilton, W. D. [1996]: *Narrow Roads of Gene Land Volume 1*, New York: W. H. Freeman.

[34] Heise, D. [1975]: *Causal Analysis*, New York: John Wiley.

[35] Heisler, I. L. and Damuth, J. [1987]: 'A Method for Analyzing Selection in Hierarchically Structured Populations', *American Naturalist*, **130**, pp. 582–602.

[36] Hölldobler, B. and Wilson, E. O. [2009]: *The Superorganism: the Beauty, Elegance and Strangeness of Insect Societies*, London: W.W. Norton.

[37] Kerr, B. and Godfrey-Smith, P. [2002]: 'On Individualist and Multi-Level Perspectives on Selection in Structured Populations', *Biology and Philosophy*, **17**, pp. 477–517.

[38] Kim, J. [1998]: *Mind in a Physical World*, Cambridge, MA: MIT Press.

[39] Lehmann, L., Keller, L., West, S. and Roze, D. [2007]: 'Group Selection and Kin Selection: Two Concepts But One Process', *Proceedings of the National Academy of Sciences USA*, **104**, pp. 6736–39.

[40] Marshall, J. A. [2011]: 'Group Selection and Kin Selection: Formally Equivalent Approaches', *Trends in Ecology and Evolution*, **26**, pp. 325–32.

[41] McElreath, R. and Boyd, R. [2007]: *Mathematical Models of Social Evolution*, Chicago: University of Chicago Press.

[42] Maynard Smith, J. [1976]: 'Group Selection', *Quarterly Review of Biology*, **61**, pp. 277–83.

[43] Maynard Smith, J. and Szathmary, E. [1995]: *The Major Transitions in Evolution*, Oxford: Oxford University Press.

[44] Michod, R. E. [1999]: *Darwinian Dynamics: Evolutionary Transitions in Fitness and Individuality*, Princeton, NJ: Princeton University Press.

[45] Michod, R. E. [2005]: 'On the Transfer of Fitness from the Cell to the Multicellular Organism', *Biology and Philosophy*, **20**, pp. 967–87.

[46] Michod, R. E. [2006]: 'The Group Covariance Effect and Fitness Trade-offs during Evolutionary Transitions', *Proceedings of the National Academy of Sciences USA*, **103**, pp. 9113–17.

[47] Michod, R. E. [2007]: 'Evolution of Individuality During the Transition From Unicellular to Multicellular Life', *Proceedings of the National Academy of Sciences USA*, **104**, pp. 8613–18.

[48] Michod, R. E. and Roze, D. [1999]: 'Cooperation and Conflict in the Evolution of Individuality. III. Transitions in the Unit of Fitness', in C. L. Nehaniv (*ed*), 1999, *Mathematical and Computational Biology: Computational Morphogenesis, Hierarchical Complexity, and Digital Evolution*. Lectures on Mathematics in the Life Sciences, vol. 26. American Mathematical Society, pp. 47–91.

[49] Michod, R. E., Viossat, Y., Solari, C. A., Nedelcu, A. M. and Hurrand, M. [2006]: 'Life History Evolution and the Origin of Multicellularity', *Journal of Theoretical Biology*, **239**, pp. 257–72.

[50] Nowak, M. A., Tarnita, C. E. and Wilson, E. O. [2010]: 'The Evolution of Eusociality', *Nature* **466**, pp. 1057–62.

[51] Nowak, M. A., Tarnita, C. E. and Wilson, E. O. [2011]: 'Nowak et al. Eeply', *Nature* **471**, pp. E9.

[52] Nunney, L. [1985]: 'Group Selection, Altruism, and Structured-Deme Models', *American Naturalist*, **126**, pp. 212–30.

[53] Okasha, S. [2004a]: 'Multi-Level Selection, Covariance and Contextual Analysis', *British Journal for the Philosophy of Science*, **55**, pp. 481–504.

[54] Okasha, S. [2004b]: 'Multi-Level Selection and the Partitioning of Covariance: a Comparison of Three Approaches', *Evolution*, **58**, pp. 486–94.

[55] Okasha, S. [2004c]: 'The Averaging Fallacy and the Levels of Selection', *Biology and Philosophy*, **19**, pp. 167–84.

[56] Okasha, S. [2005]: 'Maynard Smith on the Levels of Selection Question', *Biology and Philosophy*, **20**, pp. 989–1010.

[57] Okasha, S. [2006]: *Evolution and the Levels of Selection*, Oxford: Oxford University Press.

[58] Okasha, S. [2008]: 'Biological Altruism', in E. Zalta (*ed.*), *Stanford Encyclopedia of Philosophy*, <plato.stanford.edu>

[59] Okasha, S. [2009]: 'Individuals, Groups, Fitness and Utility', *Biology and Philosophy*, **24**, pp. 561–84.

[60] Okasha, S. [2011]: 'Reply to Sober and Waters', *Philosophy and Phenomenological Research*, **82**, pp. 241–48.

[61] Okasha, S. and Paternotte, C. [2011]: 'Group Adaptation, Formal Darwinism and Contextual Analysis', *Journal of Evolutionary Biology*, **25**, pp. 1127–39.

[62] Pearl, J. [2000]: *Causality*, Cambridge: Cambridge University Press.

[63] Price, G. R. [1970]: 'Selection and Covariance', *Nature*, **227**, pp. 520–21.

[64] Price, G. R. [1972]: 'Extension of Covariance Selection Mathematics', *Annals of Human Genetics*, **35**, pp. 485–90.

[65] Queller, D. C. [1992]: 'Quantitative Genetics, Inclusive Fitness and Group Selection', *American Naturalist*, **139**, pp. 540–58.

[66] Queller, D. C. [1997]: 'Cooperators Since Life Began', *Quarterly Review of Biology*, **72**, pp. 184–8.

[67] Queller, D. C. [2011]: 'Expanded Social Fitness and Hamilton's Rule for Kin, Kith and Kind', *Proceedings of the National Academy of Sciences USA*, **108**, pp. 10792–99.

[68] Ratnieks, F. L. W. and Reeve, H. K. [1992]: 'Conflict in Single-Queen Hymenopteran Societies: the Structure of Conflict and Processes that Reduce Conflict in Advanced Eusocial Species', *Journal of Theoretical Biology*, **158**, pp. 33–65.

[69] Roberston, A. [1966]: 'A Mathematical Model of the Culling Process in Dairy Cattle', *Animal Production*, **8**, pp. 95–108.

[70] Rousset, F. [2004]: *Genetic Structure and Selection in Subdivided Populations*, Princeton, NJ: Princeton University Press.

[71] Scheines, R. [2006]: 'The Similarity of Causal Inference in Experimental and Non-Experimental Studies', *Philosophy of Science*, **72**, *(Proceedings)*, S927–S940.

[72] Searle, J. [1992]: *The Re-discovery of the Mind*, Cambridge, MA: MIT Press.

[73] Seeley, T. D. [1997]: 'Honey Bee Colonies are Group-Level Adaptive Units', *American Naturalist*, **150**, pp. S22–S41.

[74] Shipley, B. [2000]: *Cause and Correlation in Biology*, Cambridge: Cambridge University Press.

[75] Sober, E. [1984]: *The Nature of Selection*, Chicago: University of Chicago Press.

[76] Sober, E. [2011]: 'Realism, Conventionalism, and Causal Decomposition in Units of Selection: Reflections on Samir Okasha's *Evolution and the Levels of Selection*', *Philosophy and Phenomenological Research*, **82**, pp. 221–231.

[77] Sober, E. and Lewontin, R. [1982]: 'Artifact, Cause and Genic Selection', *Philosophy of Science*, **49**, pp. 157–80.

[78] Sober, E. and Wilson, D. S. [1998]: *Unto Others: The Evolution and Psychology of Unselfish Behaviour*, Cambridge, MA: Harvard University Press.

[79] Spirtes, P., Glymour, C. and Scheines, R. [2001]: *Causation, Prediction and Search, 2nd edition*, Cambridge, MA: MIT Press.

[80] Taylor, P. D. and Frank, S. A. [1996]: 'How To Make a Kin Selection Model', *Journal of Theoretical Biology*, **86**, pp. 409–19.

[81] Taylor, P. D., Wild, G. and Gardner, A. [2007]: 'Direct Fitness versus Inclusive Fitness: How Shall we Model Kin Selection?', *Journal of Evolutionary Biology*, **20**, pp. 301–9.

[82] Traulsen, A. [2010]: 'Mathematics of Kin and Group Selection: Formally Equivalent Approaches?', *Evolution*, **64**, pp. 316–23.

[83] van Veelen, M. [2009]: 'Group Selection, Kin Selection, Altruism and Cooperation: When Inclusive Fitness Is Right and When It Can Be Wrong', *Journal of Theoretical Biology*, **259**, pp. 589–600.

[84] Wenseleers, T., Gardner, A. and Foster, K. R. [2010]: 'Social Evolution Theory: a Review of Methods and Approaches', in T. Szekely, A. J. Moore and J. Komdeur (*eds*), 2010, *Social behaviour: genes, ecology and evolution*, Cambridge: Cambridge University Press, pp. 132–58.

[85] West, S. A., Griffin, A. S., and Gardner, A. [2007]: 'Social Semantics: Altruism, Cooperation, Mutualism, Strong Reciprocity and Group Selection', *Journal of Evolutionary Biology*, **20**, pp. 415–32.

[86] West, S. A., Griffin, A. S and Gardner, A. [2008]: 'Social Semantics: How Useful Has Group Selection Been?', *Journal of Evolutionary Biology*, **21**, pp. 374–85.

[87] West, S. A. and Gardner, A. [2013]: 'Adaptation and Inclusive Fitness', *Current Biology*, **23**, pp. R577–84.

[88] Wilson, D. S. [1997]: 'Altruism and Organism: Disentangling the Themes of Multi-Level Selection Theory', *American Naturalist*, **150**, pp. S122–134.

[89] Wilson, E. O. [2012]: *The Social Conquest of Earth*, New York: W.W. Norton.

[90] Williams, G. C. [1966]: *Adaptation and Natural Selection*, Princeton, NJ: Princeton University Press.